JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

PERSPECTIVE

Applications of Metabolomics in Agriculture

RICHARD A. DIXON,[†] DAVID R. GANG,[‡] ADRIAN J. CHARLTON,[§] OLIVER FIEHN,[#] HARRY A. KUIPER,[∥] TRACEY L. REYNOLDS,[⊥] RONALD S. TJEERDEMA,[⊗] ELIZABETH H. JEFFERY,[△] J. BRUCE GERMAN,[×] WILLIAM P. RIDLEY,*[,⊥] AND JAMES N. SEIBER[◇]

Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, Oklahoma 73401; Department of Plant Sciences and Bio5 Institute, 303 Forbes Building, University of Arizona, Tucson, Arizona 85721-0036; Central Science Laboratory, Department for Environment, Food and Rural Affairs, Sand Hutton, York YO41 1LZ, United Kingdom; Genome Center, University of California—Davis, Health Science Drive, Davis, California 95616; RIKILT Institute of Food Safety, Wageningen University and Research Centre, Bornsesteeg 45, 6708 PD Wageningen, The Netherlands; Monsanto Company, 800 North Lindbergh Boulevard, St. Louis, Missouri 63167; Department of Environmental Toxicology, University of California—Davis, One Shields Avenue, Davis, California 95616-8588; Department of Food Science and Human Nutrition, University of Illinois, 905 South Goodwin Avenue, Urbana, Illinois 61081; Department of Food Science and Technology, University of California—Davis, Cruess Hall, One Shields Avenue, Davis, California 95616-8598; and Western Regional Research Center, U.S. Department of Agriculture, 800 Buchanan Street, Albany, California 94710

Biological systems are exceedingly complex. The unraveling of the genome in plants and humans revealed fewer than the anticipated number of genes. Therefore, other processes such as the regulation of gene expression, the action of gene products, and the metabolic networks resulting from catalytic proteins must make fundamental contributions to the remarkable diversity inherent in living systems. Metabolomics is a relatively new approach aimed at improved understanding of these metabolic networks and the subsequent biochemical composition of plants and other biological organisms. Analytical tools within metabolomics including mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy can profile the impact of time, stress, nutritional status, and environmental perturbation on hundreds of metabolites simultaneously resulting in massive, complex data sets. This information, in combination with transcriptomics and proteomics, has the potential to generate a more complete picture of the composition of food and feed products, to optimize crop trait development, and to enhance diet and health. Selected presentations from an American Chemical Society symposium held in March 2005 have been assembled to highlight the emerging application of metabolomics in agriculture.

KEYWORDS: Metabolomics; mass spectrometry; nuclear magnetic resonance; metabolic pathways; food; environmental safety; nutrition

INTRODUCTION

Beginning with Gregor Mendel (1822−1884), who described the inheritance of phenotypic traits in peas (*Pisum sativum*), to the classic publication, in 1953, of a proposed structure for DNA by James D. Watson and Francis Crick (*1*), critical technical and conceptual breakthroughs in our understanding of the basic biochemical, genetic and physiological processes of living systems have occurred over the past 100 years (*2*). In 2000, 47 years after Watson and Crick's paper, the full sequence of the *Arabidopsis thaliana* genome was published (*3*) followed in 2001 by publication of the human genome (*4*).

Although the complete DNA sequence of *Arabidopsis thaliana* is known, only a fraction of that sequence has been functionally characterized (*5*). Much remains to be discovered and understood before the molecular phenotype of *Arabidopsis*

* Author to whom correspondence should be addressed [telephone (314) 694-8441; fax (314) 694-8562; e-mail william.p.ridley@monsanto.com].
[†] Samuel Roberts Noble Foundation.
[‡] University of Arizona.
[§] Central Science Laboratory.
[#] Genome Center, University of California.
[∥] Wageningen University.
[⊥] Monstanto Co.
[⊗] Department of Environmental Toxicology, University of California.
[△] University of Illinois.
[×] Department of Food Science and Technology, University of California.
[◇] U.S. Department of Agriculture.

10.1021/jf061218t CCC: $33.50
Published on Web 11/04/2006

Table 1. Metabolomics Initiatives at the Noble Foundation CPNPMR

program	rationale and progress	ref
biotic and abiotic elcitors in *Medicago truncatula* cell cultures	combined with the results of DNA microarray analyses, metabolomics has helped to reveal novel pathways of isoflavonoid and triterpene secondary metabolism in addition to providing new information on pathway regulation by transcription factors and control of vacuolar storage of secondary metabolites	*13–17*
Medicago truncatula natural variants ("ecotypes")	the goal is to determine natural quantitative and qualitative variations in secondary metabolites that can be related to their genetic bases, thus facilitating pathway gene discovery	
gene annotation in *Medicago truncatula*	integrating metabolite and transcript profiling to complement genetic approaches for assigning function to the many genes that encode enzymes that "decorate" or otherwise modify natural products such glycosyltransferases, *O*-methyltransferases, acyltransferases, and cytochrome P450s	*15*
lignin biosynthesis	studies on metabolite accumulation in transgenic alfalfa lines down-regulated in a variety of enzymatic steps in the lignin pathway have revealed the regulatory architecture of the pathway and the specific sites of involvement of the candidate biosynthetic genes	*18*
gene function in transgenic legumes	metabolic profiling may serve as a powerful tool for addressing "substantial equivalence" between transgenic and nongenetically modified plants; profiling of transgenic alfalfa plants, coupled with DNA microarray analysis, has shown that genetic introduction of the isoflavone phytoestrogen genistein into alfalfa leaves is "neutral" in terms of activation of additional pathways	*19*
metabolome of plant glandular trichomes	metabolite profiles for trichomes from plants such as tobacco, tomato, and hops are being generated; these data will be integrated with EST sequence data from cDNA libraries representing trichome-expressed transcripts to support discovery of the genes involved in the biosynthesis of the trichome bioactives	*20*

can be fully described. Approaches to associating a gene with function have typically focused on quantitative and qualitative analyses of the various gene expression products—mRNA, proteins, and small molecule metabolites. Assignment of gene function is essential because it is the prerequisite for a more complete understanding of a cell, tissue, or biological organism.

Metabolomics is a promising new approach aimed at facilitating an improved understanding of the dynamic biochemical composition within living systems. This knowledge will prove to be fundamental to systems biology approaches, which attempt to synergistically integrate DNA, RNA, protein, and metabolite analyses along with phenotypic, morphological, clinical, and other biological data to provide a more holistic overview of a living system. It may also have potential in optimizing trait development in agricultural products and in biorefining. As with transcriptomics and proteomics, analytical tools within metabolomics can yield massive data sets. Data acquisition instruments typically used, such as NMR or MS, can rapidly profile the impact of time, stress, nutritional status, and environmental perturbation on hundreds of metabolites simultaneously. This can potentially generate a more complete picture of composition than traditional plant biochemistry and natural products approaches.

Agricultural crops are increasingly viewed as a source or starting point for a plant-based economy, potential input to a biorefinery in which all parts of the plant are processed and used to yield (1) food, both traditional and those with enhanced nutritional, safety, stability, processability and other desirable characteristics to meet current and projected consumer needs worldwide; (2) industrial products, including polymers, fibers, latex, industrial oils, and packaging materials, as well as basic chemical building blocks; and (3) fuels, such as hydrogen, methane, ethanol, and biodiesel (*6*−*9*).

The Agrochemical Division of the American Chemical Society sponsored a symposium at the ACS National Meeting, March 16, 2005 in San Diego, CA, titled "Applications of Metabolomics in Agriculture". The symposium, organized by William Ridley of Monsanto Co. and James Seiber of the Western Regional Research Center, USDA/ARS, addressed methods for the rapid detection, identification, and quantification of small molecules and metabolites within a sample and the potential relevance of such results. The purpose of the symposium was to assemble key international research scientists to provide an appreciation of the technical challenges associated with metabolomics, its current application in agriculture in terms of plant biochemistry and food and environmental safety, and its potential to be used as a tool to improve nutrition, diet, and health. The following sections summarize selected presentations and describe their significance. This paper does not attempt to comprehensively cover all applications of metabolomics in agriculture but rather to highlight areas of special interest to the participants in the symposium.

UNDERSTANDING METABOLIC PATHWAYS THROUGH METABOLITE ANALYSES

Metabolites are often simply viewed as one of the end-products of gene expression and protein activity. It is increasingly understood that metabolites themselves modulate macromolecular processes through, for example, feedback inhibition and as signaling molecules. Metabolomic studies are therefore intended to provide an integrated view of the functional status of an organism. Richard Dixon has partnered with colleagues at the Noble Foundation, Lloyd Sumner and Xiaoqiang Wang, to integrate a metabolomic platform with natural product expertise and structural biology and genomics capabilities, a program coordinated through the Center for Plant Natural Product and Metabolomic Research (CPNPMR). The CPNPMR seeks to utilize this "systems biology" platform to understand the regulation of natural product syntheses, identify novel pathways of isoflavonoid and triterpene metabolism, and provide new information on pathway regulation by transcription factors (*10*−*12*). Ongoing initiatives are presented in **Table 1**.

As a new and maturing science, metabolomics is confronted by immense challenges related to data acquisition (*20*). Metabolites represent a diverse range of structures, physicochemical properties, stabilities, and abundances. A key consideration in effective metabolomic pursuits is, therefore, the establishment of an optimal balance between quantitative accuracy and the range of metabolites measured. To address issues related to metabolome data acquisition, Lloyd Sumner of the Noble Foundation has designed a strategy that utilizes sequential or selective extraction of plant tissues followed by parallel data acquisition on each extract. This parallel analysis is designed

to provide a comprehensive view of the metabolome and utilize an arsenal of analytical techniques including gas chromatography (GC)−mass spectrometry (MS), high-performance liquid chromatography (HPLC)-MS, and capillary electrophoresis (CE)-MS.

Sequential and selective extraction is intended to generate discrete manageable classes of physicochemically related compounds, thereby optimizing the accuracy of data acquisition by facilitating the use of parallel analytical profiling techniques specific for a given metabolite class. Methods are currently being employed or developed for most primary and secondary plant metabolites including soluble sugars, sugar phosphates, complex carbohydrates, amino acids, organic acids, alcohols, lipids, sterols, phenylpropanoids (including flavonoids), lignins, and triterpene saponins.

A major challenge for this type of approach is the integration of multiple data sets reporting not only metabolite information but also detailed transcript and proteomic profiling data. In collaboration with the group of Pedro Mendes at the Virginia Bioinformatics Institute, Dixon, Sumner, and colleagues have populated and developed DOME (*d*atabase for *OME*s) to store and allow for interrogation of functional genomics data including, but not limited to, DNA microarray data, protein fragment mass spectral data from 2D gel separations, and metabolite levels determined by MS after separation by GC, LC, or CE (*21*). The DOME database will become publicly available in 2006. Additional databases and programs that allow integration of metabolite with transcript data are AraCyc (http://arabidopsis.org/tools/aracyc/) (*22*), MAPMAN (*23*), and KaPPA-View (http://kpv.kazusa.or.jp/kappa-view/) (*24*).

In a manner analogous to that pursued by the CPNPMR, David Gang and colleagues, at the University of Arizona, have sought to elucidate the biosynthetic pathways that produce novel and important specialized metabolites in aromatic plants, to uncover the mechanisms responsible for the evolution of these pathways, and to understand the function of a given natural product within the biology and physiology of a given plant species (*25*−*28*). This group has conducted metabolic profiling experiments using selected lines and tissues of ginger (*Zingiber officinale* Rosc.), turmeric (*Curcuma longa* L.), and sweet basil (*Ocimum basilium* L.) and described the complexity of metabolic pathways that can exist in such plants. In addition, aromatic plant Expressed Sequence Tag (EST) databases from ginger and turmeric tissues and from sweet basil glandular trichomes have been assembled and used to identify several genes involved in the production of these specialized metabolites in these plants. The expression profiles of genes in various tissues and at various stages of development coupled with LC-MS and GC-MS metabolic profiling are now being pursued to characterize genes involved in the production and regulation of specialized metabolites (*29*, *30*). The long-term goal is to use this knowledge for rational custom-designed breeding by classical methods as well as the application of genetic engineering techniques to improve and develop new aromatic plants. Having such databases available to the public is essential to move progress in -omics studies forward. The EST databases are available through the Gang Laboratory Web page at http://ag.arizona.edu/research/ganglab/basilESTs.htm and at http://ag.arizona.edu/research/ganglab/ArREST.htm.

The metabolomic approaches pursued by the CPNPMR and by Gang's laboratory primarily exploit MS as the data acquisition technology of preference. This is primarily attributable to its greater sensitivity and dynamic range of metabolite concentration when compared to other techniques such as nuclear magnetic resonance (NMR) spectroscopy. However, the uniformity, reproducibility, and ubiquity of the NMR response coupled with the requirement for limited sample preparation make NMR the ideal tool for broad-range profiling of abundant metabolites and for metabolite fingerprinting of extensive sample collections. Constitutive levels of metabolites readily respond to perturbations in environmental conditions, including breeding processes as well as direct molecular bioengineering approaches. NMR profiling, typically in conjunction with multivariate analyses, is particularly adept at identifying such changes. In other words, NMR-based metabolite profiling is well suited to monitor and quantify the degree of metabolic impact induced by genetics, environment, or bioengineering.

Adrian Charlton, at Central Science Laboratory, York, U.K., and his colleagues have applied NMR analysis to pea (*Pisum sativum*) extracts to define the impact of environment and genetic diversity on baseline metabolite profiles (*31*, *32*). The plants, as discussed at the symposium, included 20 diverse lines of *Pisum*, 20 recombinant inbred lines (RILs) derived from a wide cross (between *P. sativum* cv. Ethiopia and *P. sativum* cv. Cennia), four independent transgenic lines derived from *Agrobacterium*-mediated transformation of *P. sativum* (cv. Puget) using a construct composed of a trypsin inhibitor (TI) gene promoter-*GUS* fusion and a *bar* selectable marker gene, and control lines that were azygous segregants, identified at T_2 or T_3 generations for every transformed line. This allowed eight data groupings (identified in parentheses) and defined as follows:

(i) RILs grown in 2003 (RI-03)

(ii) RILs grown in 2004 (RI-04)

(iii) cv. Puget plants sown at monthly intervals (Mnth) under glasshouse conditions in an experiment designed to investigate the effect of growing season on metabolite profiles

(iv) cv. Puget plants grown in a range of environments (Envi) including a controlled environment room, glasshouse, polytunnel, and outdoor plots

(v) diverse germplasm lines harvested in 2003 (GP03)

(vi) drought-stressed transgenic and azygous plants (Drought) grown in one season

(vii) well-watered transgenic plants (W/T) grown in one season

(viii) well-watered azygous plants (W/N) grown in one season

Leaf samples were harvested from all plants at the onset of flowering, and seeds were harvested at maturity, following desiccation. NMR analysis on these samples followed a previously established protocol (*32*). The NMR profiles were subjected to statistical analysis using a number of approaches, including principal components analysis (PCA), partial least-squares discriminant analysis (PLS-DA), analysis of variance (ANOVA), and Student's *t* test. Spectra were subjected to statistical analysis to determine the regions of the NMR spectra that exhibited the greatest variation and also to identify profiles that could be correlated with genotype, environment, or transgenesis. PCA identified the watering regimen as the factor that affected leaf composition most significantly. Clear separation between the well-watered plants and those that had been subjected to drought was seen when the first two principal component scores were plotted. It was also apparent that the genotype of the plants significantly affected the metabolome. Many of the individual genotypes from the diverse germplasm formed tight and distinct data clusters (*33*) when subjected to PCA. Using PLS-DA, it was possible to classify the leaf extracts as derived from transgenic, or azygous, plants, and this classification was more successful when the plants were well-watered than when they were drought-stressed. It was not

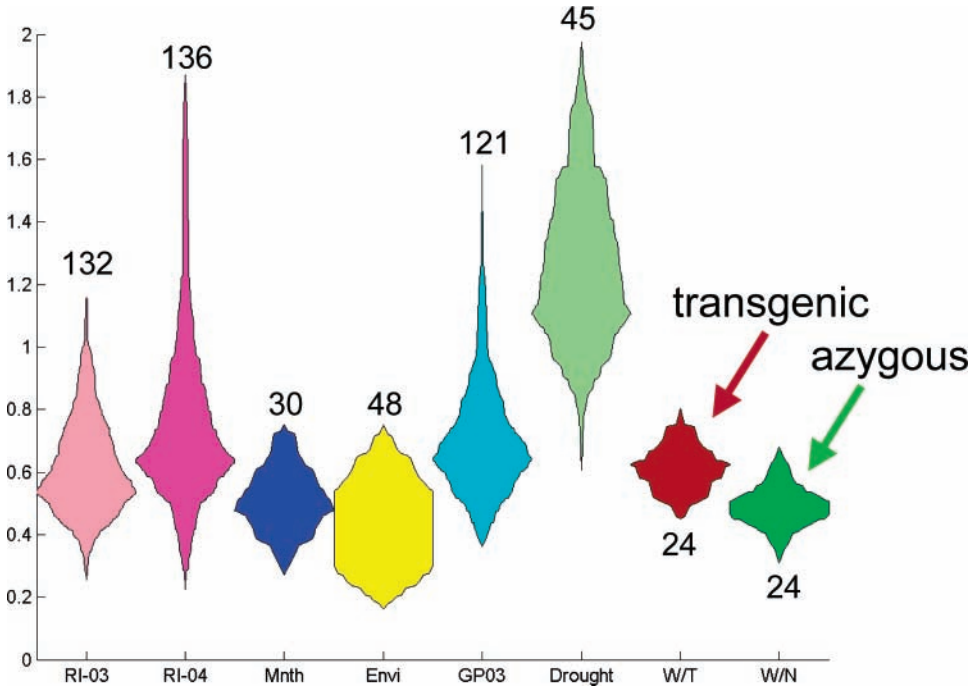

Figure 1. Range of NMR intensities observed for a metabolite resonating at 2.33 ppm. The range of metabolite concentrations for each treatment is shown by the height of the colored plot for each group. The width of the colored areas represents the proportion of the total number of observations (in each group) having a particular NMR intensity. The number of plants contributing to each group is indicated above or below the group. The group names are defined in the text.

possible to separate the transgenic plants from the azygous controls when both the drought-stressed and well-watered plants were treated as one group. This strongly suggests that any impact on the metabolome that is conferred by the function of the transgene may be masked by larger environmental effects, such as drought.

Figure 1 shows the NMR intensities at approximately 2.33 ppm, determined for the eight groups. The results of the *t* test performed on the well-watered transgenic (W/T) and azygous control (W/N) groups show that the metabolite that gave rise to this resonance is present at an elevated level in the transgenic plants ($p < 0.0001$). It is also clear that the wide range of genetic diversity for this metabolite, observed in both the RILs (RI-03, RI-04) and diverse germplasm (GP03), exceeds the elevated range determined for the transgenic plants. In the case of this particular metabolite, the difference between the transgenic and azygous control groups was statistically the most significant single-point difference in the pea leaf metabolome, as observed by NMR. However, when the range of concentrations of this metabolite, as affected by environmental factors and genotype, is considered, it is clear that the relevance of this significant difference in the context of genetically modified (GM) safety evaluation is minimal.

In agrobiotechnology, this complex interaction of genes and a multitude of environmental conditions leads to large genotype × environment (G × E) experimental designs, which easily generate hundreds to thousands of individual samples. Most published literature on plant metabolomics describes studies of fewer than 100 samples. However, field trials and comprehensive G × E plots call for sample sizes that will require analysis over weeks and years, in order to store, disseminate, and query metabolic phenotypes that may be needed to guide the generation of further crop genotypes or even to support regulatory purposes. Therefore, validated sample preparation protocols and monitoring of quality control are required to ensure the long-term robustness of data acquisition.

Recently, an overall data model has been proposed for capturing the different components of plant metabolomic experiments aimed at comparability and reusability of data: an Architecture for Metabolomics, ArMet (*34*). This architecture is based on nine modules, among them the experimental design modules "BioSource" and "Growth", but also containing modules reporting on sample preparation, data collection, data analysis, and data treatment. The ArMet report strongly emphasized that actual implementations may differ from the nine-module structure. Consequently, Oliver Fiehn and colleagues at the University of California—Davis have developed a database implementation that more strongly focuses on the description of the underlying biological study, which was suggested to generally follow a G × E experimental design, but which differed from the ArMet schema in two important features: (a) each given subdescription of the BioSource (e.g., inclusion of different genotypes and different organs in a study) will define statistically different classes and (b) growth conditions are now subdivided in hierarchy levels of increasing detail, from general growth conditions that apply to all plants within a study to different treatments (such as abiotic stress or soil types) to the duration and intensity of such treatments. Each G × E design defines a class or a statistically independent group that is used for randomizing the injection sequence in the analytical laboratory (which may span several months, if the study design is large) but also serves as a template for randomizing analysis sequences in the laboratory.

This database schema has been implemented as *SetupX* system since June 2005 at the University of California—Davis Genome Center metabolomics laboratories (http://fiehnlab.ucdavis.edu/db/setupx). It is seamlessly integrated with a multilayered data processing algorithm called *BinBase* to filter and annotate metabolite profiling results (*35*), which are then exported to the user for statistical investigation. As a test case for the implementation of the "study design database" *SetupX* and its integration with the subsequent raw data processing

database *BinBase*, over 1300 potato tuber chromatograms were assigned to their corresponding 48 experimental design classes, consisting of 12 genetically different plant lines grown in 4 slightly different soil plots. Six GM potato lines were selected, three lines based on the level of gene overexpression of a sucrose:sucrosyltransferase (SST) and three lines based on overexpression of a fructose:fructosyltransferase in addition to the first genetic modification (SSTFFT). These six GM lines were compared to six conventional cultivars. Potatoes were planted in four blocks of different soil types and grown throughout the spring and summer of 2003, without any further treatment or growth-related experimental design. A total of 48 classes (12 genotypes and 4 soil blocks) were compared.

The *SetupX* system is linked with the NCBI taxonomy database, so that the metabolomic database can also be queried for results of species that are related in taxonomic terms. Although the NCBI taxonomy database is not an authoritative primary source for taxonomic or phylogenetic information, it presents a very comprehensive and continuously updated source of information of species and their phylogenetic relationships. Consequently, it may serve general metabolomic annotations better than databases that focus on specific kingdoms or species. The inclusion of taxonomic information even at the level of data entry further serves convenience in use because either systematic names or synonyms can be used by researchers. For the case of a "potato" field trial, a species definition such as "potato" is automatically replaced by "*Solanum tuberosum*", and additional subspecies may be selected if available (*Solanum tuberosum* subsp. *andigena*). However, a genus entry alone is insufficient such as for the case of "*Arabidopsis*", and users are subsequently asked to specify a species. An extension of the *SetupX* database is planned by linking BioSource entries to content definitions given by the plant ontology consortium (http://www.plantontology.org).

After the physical objects (genotypes and organs, tissues or cell types) have been defined in BioSource, "growth" description is requested where growing conditions that were overall identical for all samples in the study, such as climatic data for the growing seasons, soil type, or plant developmental stage at harvest, are entered. If a G × E experiment was carried out at different sites, each site might have a different "growth" description leading to different classes. Similarly, classes are generated by different "treatment" designs, such as different types of fertilizers and/or different degrees of fertilization. It is important to note that the "sample preparation" module consists of different layers that call for extraction and derivatization Standard Operating Procedures (SOPs), but may also generate classes if different protocols were used on the same type of plant material as part of method development and validation. However, no module differentiation is given for the actual analytical and data processing steps, because these are defined in the metabolomics laboratories and the associated *BinBase* database and cannot be altered or defined by external users. The use of the *SetupX* system should facilitate the collection and analysis of large-scale field studies designed to evaluate G × E interactions and the associated variability in the metabolome.

This schema of combining study design with standard operating procedures and quality control has been implemented in a metabolomic database and demonstrated for a study on "substantial equivalence" of genetically modified potatoes compared to classical cultivars (*36*). The concept of substantial equivalence is a key step in the safety assessment process and is used to identify similarities and differences between the genetically modified food and its closest traditional counterpart with a history of safe use (*37*, *38*). Application of the concept is not a safety assessment in itself but a starting point to structure the safety assessment of the genetically modified food relative to its traditional counterpart.

The World Health Organization and the United Nations Food and Agriculture Organization acknowledged in May 2000 (*38*) that some aspects in the safety assessment process for novel foods could be refined. They indicated that profiling techniques following validation could be considered as a useful supplement to current practices for compositional analysis of novel foods. As discussed above, profiling crops using a diverse set of analytical and bioinformatics tools with a broad range of detection capabilities has the potential, when fully validated, to enhance the understanding of crop composition. These issues will be more fully explored in the next section.

FOOD AND ENVIRONMENTAL SAFETY

The European Thematic Network on the Safety Assessment of Genetically Modified Foods (ENTRANSFOOD) is funded by the European Commission as part of the 5th Framework program (www.entransfood.com). The Network has developed a detailed and integrated approach for the safety assessment of foods derived from GM crops by assessing the characteristics of (i) the parent crop, (ii) the donor, transgenes, and the genetic modification process, (iii) the expressed gene product(s), and (iv) the new GM crop/food. Harry Kuiper, at RIKILT Institute of Food Safety, Wageningen University and Research Centre, The Netherlands, a leader of this initiative, outlined the objectives and research programs of the participating groups (*39*). His presentation focused on the detection of potential unintended alterations in the composition of GM food crops. It should also be noted that unintended effects on the concentrations of plant constituents often occur in classical plant breeding through recombination or mutagenic events. Reports of unintended effects in conventionally bred plants have appeared: for instance, high contents of furanocoumarins in celery and glycoalkaloids in potatoes or high cucurbitacin levels in squash (*40*). In practice, extensive back-crossing during classical breeding is aimed at the removal of unintended effects, thus preventing serious food safety problems.

Detection of unintended effects in GM food crops relies on targeted approaches, that is, comparative determination in GM and non-GM products of levels of selected macro- and micronutrients, antinutrients, and known toxins. Limitations of this approach lie in the still fragmented knowledge of plant metabolic pathways and the "biased" selection of compounds. To increase the probability of detecting unintended effects, Kuiper suggests that nontargeted approaches, using comprehensive profiling or -omics techniques, should be further developed and validated (*40*, *41*). These profiling methods may be of particular relevance in the case of GM food crops with altered nutritional and/or health beneficial properties obtained through multiple gene insertions and modification of specific pathways. The approaches identified by Kuiper included gene expression analysis, proteomics (2D gel electrophoresis and MALDI-TOF-MS), and metabolic profiling (LC-MS/MS and LC-NMR).

Microarrays of ESTs from green and red-ripe conventionally bred tomatoes hybridized with mRNA extracted from the tomatoes in the different ripening stages showed distinct hybridization patterns. This suggested that potential differences in levels of compounds such as tomatins, which alter during ripening, may be traced if changed as a result of genetic modification. Proteomic analyses of these tomatoes also indicated distinct differences in expression profiles. This suggests

large variation in gene expression and protein profiles during normal tomato development.

Results were also presented of gene expression analyses of GM tomatoes (cv. Ailsa Craig), which had been altered in carotenoid and flavonoid content. This study indicated that the intended changes in the levels of these compounds could be identified by transcriptomic profiling. ^{1}H NMR-based metabolic profiling of the GM Ailsa Craig tomatoes and its azygous control indicated that GM Ailsa Craig tomatoes contained 5−8 times higher concentrations of cyclopropyl-sterol (cycloartenol) compared to controls. This change may be considered as an intended effect.

Kuiper concluded that the results from microarray technology, proteomics, and metabolite profiling may be successfully applied to screen for intended and unintended side effects of GM foods. However, before these techniques can routinely be used for safety assessment, much additional work needs to be done with respect to the development of uniform sampling and extraction procedures, instrument methodology optimization, multivariate data analysis strategies, and databases containing profiling data from both microarray and compositional analysis under diverse developmental stages and environmental conditions. Currently, efforts are underway initiated by the Metabolomics Society to develop and implement minimum reporting standards and current best practice protocols in the "Metabolomics Standards Initiative" (MSI) (http://msi-workgroups.sourceforge.net/).

As implied above, new plant products developed through biotechnology are required to undergo extensive safety and nutritional assessment prior to marketing as human and animal feed. A critical component of this is the analysis of key nutrients and antinutrients present in the plant. Guidance regarding the choice of analytes to monitor is provided from organizations such as the Organization for Economic Cooperation and Development (OECD) that publish consensus documents on compositional considerations for new plants (*42*). These documents provide scientifically based guidelines that are acceptable to national regulatory bodies for the relevant nutrients and antinutrients to be analyzed for new crop varieties.

Compositional assessments of biotechnology-derived (biotech) crops require analyses, not only of the biotech and parental control hybrids but also of a range of conventional commercially available hybrids grown at the same locations. Analytical data from the conventional hybrids can be used to develop a tolerance interval, defined as a statistically determined range that represents a given percentage of the commercial population. This interval is utilized to place any observed analyte differences between the biotech and parental crop within a contextual framework of natural variation for that analyte (*43*, *44*).

The concept of natural variability of crop analytes is fully recognized, yet the extent of this variability is rarely considered when direct comparisons between biotech and control lines are pursued. Tracey L. Reynolds and her colleagues at Monsanto Co. recently described a study to understand baseline variability in hybrid corn grain composition (*45*). As part of this study, seven conventional corn hybrids were grown in four different locations in the European Union, and grain was harvested and analyzed for a range of metabolites using classical, quantitative methodology (*45*). These analytes included proximates (moisture, protein, fat, ash, and carbohydrate), fatty acids, amino acids, vitamins, and minerals. Of the 4935 hybrid-to-hybrid analyte comparisons possible, 40% were found to be statistically significantly different ($p < 0.05$). The largest differences were seen in the fatty acids, minerals, and protein. In these instances, the difference from the mean ranged from 0.84 to 149%. To assess the effects of environment on corn composition, comparisons were made between analytes for a single hybrid grown at all sites. Of the 1974 comparisons possible for an individual hybrid across sites, 22% (429) were found to be statistically significantly different ($p < 0.05$). The individual hybrids had statistical differences ranging from 16 to 27% of the total comparisons. The largest differences were found in minerals, specifically calcium, manganese, and copper. In summary, this study emphasized the range of natural variation in the biochemical composition of commercially available crops and the need to understand this variability when direct comparisons between a biotech and control line are made.

Over the past 15 years, Ron Tjeerdema and colleagues, at the University of California—Davis, have focused on applying NMR-based techniques to environmental safety issues. Originally, surface-probe ^{31}P NMR was used to elucidate the in vivo actions of natural stresses (e.g., hypoxia, salinity, or temperature change) or agricultural pesticides on energetics in the eggs, larvae, or adults of various aquatic invertebrates and fishes (*46*−*60*). As the action of electron transport inhibitors or oxidative phosphorylation uncouplers can be monitored in real time, in vivo NMR represents a powerful approach to assess biochemical mechanisms in intact organisms. However, limited signal sensitivity restricts its use in assessing the impacts of agrochemicals on whole organism health.

In recent years, Tjeerdema's group has begun to apply more sensitive ^{1}H NMR-based metabolomics to characterize organismal health and the metabolic perturbations that result from exposure to either natural stress factors (*61*−*63*) or hazardous chemicals (*64*−*66*) in the environment. For example, they have used this approach to determine the metabolic actions of withering syndrome, a metabolic disease caused by the pathogen *Candidatus xenohaliotis californiensis* (a gastrointestinal intracellular Rickettsiales-like prokaryote) in abalone (*Haliotis* spp.; *61*, *63*). They have also investigated the actions of agricultural pesticides (diazinon, dinoseb, and esfenvalerate; *65*, *66*) or solvents (trichloroethylene; *64*) on the developing life stages of Chinook salmon (*Oncorhynchus tshawytscha*) and Japanese medaka (*Oryzias latipes*).

Specifically, the elucidation of the actions of trichloroethylene (TCE; at 0, 8.76, 21.9, 43.8, 87.6, and 175 mg/L) on medaka embryos (*64*) and the relative sensitivities of the traditional and metabolomic endpoints were compared. Although the no-observable-effect level (NOEL) for hatching success (the most sensitive traditional endpoint) was 164 mg/L TCE, metabolic perturbations as detected by ^{1}H NMR were observed at all concentrations. Additionally, 12 metabolites that exhibited highly significant dose−response relationships were identified, suggesting a high energetic cost from TCE exposure (**Figure 2**).

Additionally, embryos were exposed to low levels of TCE and sampled on each of the 8 days of embryogenesis. Projections of two-dimensional (2D) *J*-resolved NMR spectra were obtained, and PCA revealed developmental metabolic trajectories that characterized the basal and TCE-perturbed changes in the entire NMR-visible metabolome throughout embryogenesis (**Figure 3**). Although no occurrences of mortality, gross deformity, or developmental retardation were apparent, TCE-induced metabolic perturbations were observed by the eighth day. Thus, the fundamental advantage of metabolomics for assessing chemical toxicity over morphologic endpoints is increased sensitivity and the ability to monitor hundreds of metabolites simultaneously, providing a more comprehensive assessment of whole organism health.

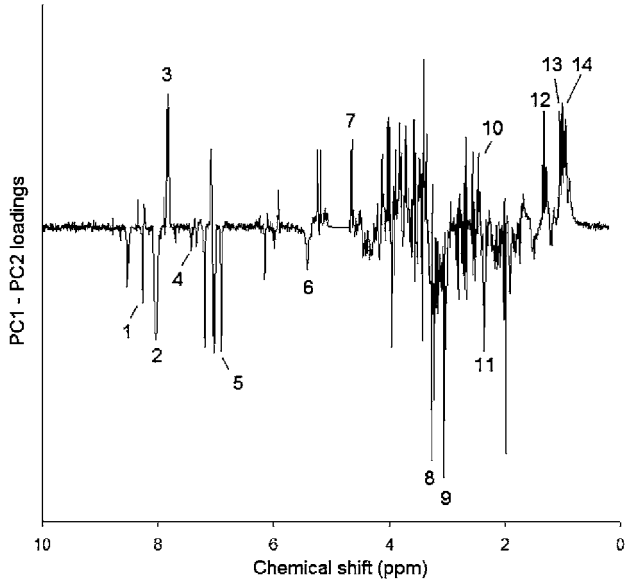

Figure 2. PCA loads plot (for the PC1–PC2 axis) from analysis of the 1D NMR spectra of medaka embryos exposed to a series of five TCE concentrations. Positive peaks correspond to metabolites that are at higher concentration in TCE-exposed embryos, whereas negative ones correspond to metabolites that are at lower concentration. Peaks: 1, ATP; 2, histidine-like compound (unassigned); 3, histidine; 4, phenylalanine; 5, tyrosine; 6, glycogen; 7, glucose; 8, glycine-betaine; 9, phosphocreatine; 10, glutamine; 11, glutamate; 12, lactate; 13, valine; 14, isoleucine.

FUTURE DIRECTIONS IN NUTRITION, DIET, AND HEALTH

The identification of dietary components that can slow or prevent chronic disease progression in humans is an exciting and active field today. Yet results from epidemiological and controlled clinical studies on bioactive food components have been inconsistent in demonstrating the type of robust protection against chronic disease progression that would be predicted from the many studies on the effects of foodstuffs in cell culture and animal models. For example, of 55 epidemiological studies that evaluated the relationship between crucifer ingestion and risk for cancer, 22 failed to report any inverse relationship (*67*). Of seven cohort studies of crucifer intake on cancer progression, none revealed any significant beneficial effect (*68*). Reasons for the failure to consistently translate efficacy in test model systems to the human situation are manifold. Major contributors include variation in levels of the putative bioactive components within the plant, the effects of food processing, and the varied response of genetically different individuals.

The impact on public health when exposure to bioactive food components drops below threshold levels is clearly important and must be addressed. The laboratory of Elizabeth Jeffery, at the University of Illinois, has sought to identify causes for the variation in efficacy observed in studies on the anticancer activity of broccoli, with a focus on sulforaphane and other anticarcinogens. Together with plant geneticists from the laboratory of John Juvik, Jeffery has identified substantial variation in the content of glucosinolates across broccoli varieties and across growing seasons. The major indolyl glucosinolate, glucobrassicin, is hydrolyzed to indol-3-carbinol, a bioactive component that may protect against breast cancer. Levels of indolyl glucosinolate have a greater dependence on environment (33% of variation) than genotype (12% of variation) (*69*). The relatively minor role of genotype in dictating glucobrassicin levels clearly emphasizes the importance of effective pre- or postharvest treatments in promoting sufficient dietary exposure to the anticancer effects of glucobrassicin.

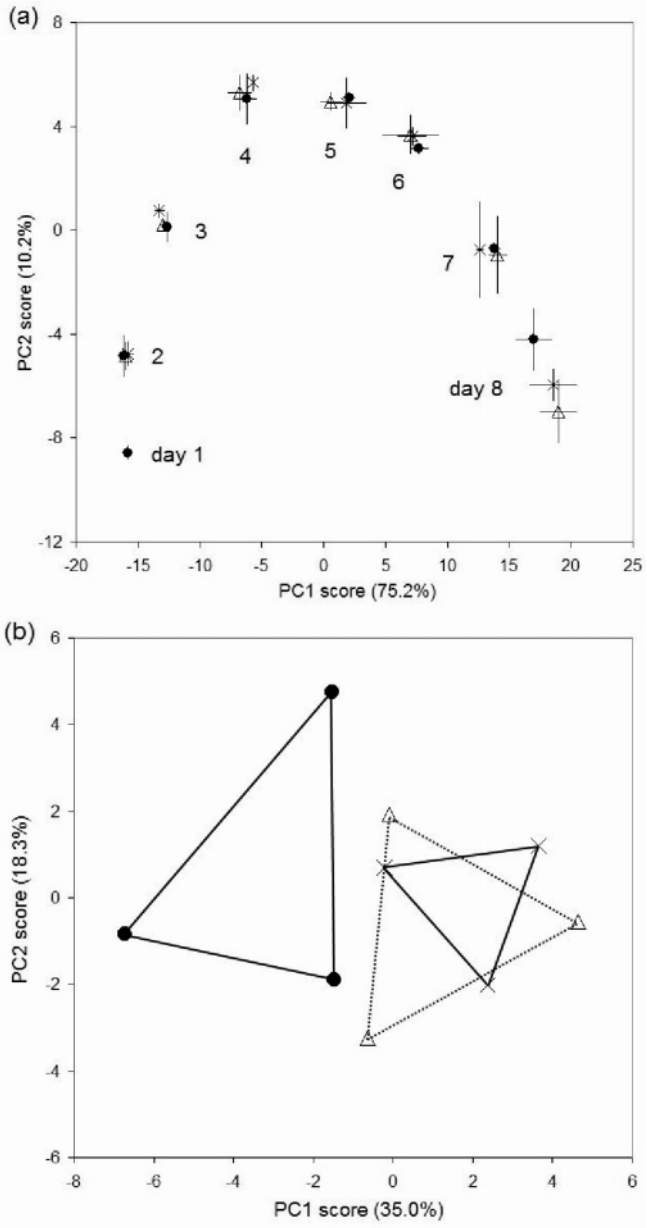

Figure 3. PCA scores plots from the analysis of the projections of 2D *J*-resolved NMR spectra of medaka embryos exposed to 0.0 (●), 0.88 (△), or 8.76 (×) mg/L TCE: (**a**) analysis includes all time points from fertilization (day 1) to immediately prior to hatching (day 8), illustrating the concept of developmental metabolic trajectories that characterize the basal and TCE-perturbed metabolic changes throughout embryogenesis; (**b**) analysis of the 8-day-old embryos only, to emphasize the metabolic perturbation induced by TCE.

On the other hand, variation in the content of aliphatic glucosinolates such as glucoraphanin, the thioglycoside precursor of the bioactive sulforaphane, is >60% dependent upon genotype (*69*). This would appear to support the development of broccoli hybrids bred to provide enhanced levels of glucoraphanin (*70*). However, the Jeffery laboratory has established that there is no correlation between glucoraphanin levels and yield of sulforaphane. When broccoli is crushed, glucoraphanin is hydrolyzed by the enzyme myrosinase. This yields not only sulforaphane (an isothiocyanate) but, to a far greater extent (80%), the inactive sulforaphane nitrile (*71*, *72*). The

Jeffery laboratory has now shown that formation of inactive nitrile is driven by a myrosinase cofactor, termed the epithiospecifier protein, which is able to complex and remove the sulfur atom (*73*). There is considerable variability in levels of the epithiospecifier protein expressed in broccoli varieties, implying that it may be under strong genetic control and that broccoli varieties may be developed in the future that lack expression of epithiospecifier protein. This, in principle, would provide a far greater yield of the bioactive sulforaphane (*73*).

Interestingly, steaming broccoli for as little as 1−2 min destroys the heat-labile epithiospecifier protein, whereas continued heating destroys myrosinase activity. Thus, overcooked broccoli provides a diet of nonhydrolyzed glucosinolates. Although these glucosinolate metabolites present no anticarcinogenic activity, gut microflora within the mammalian host are able to hydrolyze the thioglucose, thus yielding bioactive glucoraphanin. In other words, issues related to modifying levels of epithiospecifier protein may be rendered moot if broccoli were heated sufficiently during food preparation. Preliminary studies in human subjects show that although some sulforaphane metabolites are excreted in urine following the ingestion of nonhydrolyzed glucosinolates, the yield is much greater from either raw or briefly heated broccoli. Interestingly, not all Crucifereae express this epithiospecifier protein. Thus, radish (*Raphanus sativus* L.) and mustard (*Sinapis alba* L.) yield only bioactive isothiocyanates upon hydrolysis (*74*) due to the absence of epithiospecifier protein.

The studies described above indicate that the precursor to a bioactive component may not provide a high or consistent yield. This can be attributed to the formation of alternate products, the effects of processing, or even altered bioavailability. It is apparent that rather than monitoring the effects of genotype, environment, and processing on a single putative bioactive component, it may be prudent to study an entire group of biosynthetically related metabolites and/or enzymes that constitute a pathway. Such "metabolic profiling" or "targeted" metabolomics approaches may identify stable genotypes that can be used to provide plants with consistently expressed amounts of a bioactive component. This, in turn, would allow epidemiological and clinical studies to more consistently reflect potential health benefits.

Developments in the field of nutrigenomics (see www.nugo.org) are now revealing that human genotypes may dictate differing dietary needs to maintain health. However, given the range in composition and function (and choice!) of food, coordinating appropriate dietary intake with the immediate health requirements of an individual represents a formidable challenge in the 21st century (*75*). In the future, health assessments may require integrating accurate measurements of the human metabolome with accurate metabolomic assessments of foodstuffs. An overview of lipids (also referred to as "lipidomics") provides context within which to model the overall scientific challenge. Lipids are responsible for much of the beneficial sensory attributes and, thereby, preferences for particular foods. However, the overconsumption of specific classes of fats by individuals is implicated in the etiology of many human chronic and nonchronic diseases.

It has been argued that we must (i) provide a more accurate and mechanistic understanding of the nutritional impact of different dietary lipids and (ii) understand variation in humans in terms of their needs for, and responses to, those lipids (*75*). Research in Bruce German's laboratory, at the University of California—Davis, seeks, through metabolomics and genomics advances, to bring to fruition a system of individual health assessment coupled with guided dietary composition. From an agricultural perspective, this would lead to more precise estimates of benefits and risks of modulating the lipid composition of commodities and foods. From a health perspective, this knowledge will develop quantitative and predictive indices of metabolic status and how diet can influence metabolism toward a net beneficial long-term trajectory and better health. An intriguing question being asked by the German laboratory is "Precisely what diets improve not just one aspect of disease risk, but overall improved health when consuming them?" As such, German is investigating milk, the only food biomaterial designed by evolutionary pressure to explicitly nourish and improve the overall health of mammals. Bioinformatics approaches are now collating the entire family of genes responsible for lactation (the milk genome) in various mammals and to combine this with biochemical methods to deconstruct milk composition. It is proposed that this will ultimately lead to the annotation of specific milk components with the biological benefits they provide to consumers (*76*).

CONCLUSIONS

These examples show the breadth of metabolomic studies related to agriculture and food, despite the nascent stage of the field. Since this ACS symposium was held, several reviews of metabolomics and their application in plant sciences have now appeared (see, e.g., refs *77*−*79*). Some of these have addressed development in analytical technologies for metabolome data acquisition, whereas others have discussed approaches to managing and interrogating the large datasets generated in metabolomic experiments (*80*−*82*) or visualizing metabolic information (*83*). Although much progress has been demonstrated, challenges remain before the complex information generated by metabolomic research can be fully interpreted and applied to an understanding of the biological systems upon which modern agriculture is based.

ACKNOWLEDGMENT

We thank Dr. George Harrigan for reviewing the manuscript and providing insightful comments and suggestions.

LITERATURE CITED

(1) Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737−738.

(2) Ridley, W. P. Introduction to agricultural biotechnology: challenges and prospects. In *Agricultural Biotechnology: Challenges and Prospects*; Bhalgat, M. K., Ridley, W. P., Felsot, A. S., Seiber, J. N., Eds.; ACS Symposium Series 866; American Chemical Society: Washington, DC, 2004; Chapter 1, pp 3−17.

(3) The *Abrabidopsis* initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature* **2000**, *408*, 796−815.

(4) Venter, J. C., et al. The human genome. *Science* **2001**, *91*, 1304−1351.

(5) Fiehn, O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155−171.

(6) Abbas, C. A.; Cheryan, M. Biorefinery opportunities. *Appl. Biochem. Biotechnol.* **2002**, *98*, 1147.

(7) Bungay, H. Product opportunities for biomass refining. *Enzyme Microb. Technol.* **1992**, *14*, 501−507.

(8) Koutinas, A. A.; Wang, R.; Webb, C. Restructing upstream bioprocessing: technological and economic aspects for production of a generic microbial feedstock from wheat. *Biotechnol. Bioeng.* **2004**, *85*, 524−538.

(9) Realff, M. J.; Abbas, C. Industrial symbiosis: refining the biorefinery. *J. Ind. Ecol.* **2004**, *7*, 5−9.
(10) Dixon, R. A. Engineering plant natural product pathways. *Curr. Opin. Plant Biol.* **2005**, *8*, 329−336.
(11) Kutchan, T.; Dixon, R. A. Secondary metabolism: nature's chemical reservoir under deconvolution. *Curr. Opin. Plant Biol.* **2005**, *8*, 227−229.
(12) Dixon, R. A.; Sumner, L. W. Legume natural products. Understanding and manipulating complex pathways for human and animal health. *Plant Physiol.* **2003**, *131*, 878−885.
(13) Dixon, R. A.; Achnine, L.; Deavours, B. E.; Naoumkina, M. Metabolomics and gene identification in plant natural product pathways. In *Biotechnology in Agriculture and Forestry, "Plant Metabolomics"*; Saito, K., Dixon, R. A., Willmitzer, L., Eds.; Springer: Berlin, Germany, 2005; pp 243−259.
(14) Achnine, L.; Huhman, D. V.; Sumner, L. W.; Blount, J. W.; Dixon, R. A. Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula. Plant J.* **2005**, *41*, 875−887.
(15) Broeckling, C. D.; Huhman, D. V.; Farag, M.; Smith, J. T.; May, G. D.; Mendes, P.; Dixon, R. A.; Sumner, L. W. Metabolic profiling of *Medicago truncatula* cell cultures reveals effects of biotic and abiotic elicitors on metabolism. *J. Exp. Bot.* **2005**, *56*, 323−336.
(16) Suzuki, H.; Reddy, M. S. S.; Naoumkina, M.; Aziz, N.; May, G. D.; Huhman, D. V.; Sumner, L. W.; Blount, J. W.; Mendes, P.; Dixon, R. A. Methyl jasmonate and yeast elicitor induce differential genetic and metabolic reprogramming in cell suspension cultures of the model legume *Medicago truncatula. Planta* **2005**, *220*, 698−707.
(17) Chen, F.; Duran, A. L.; Blount, J. W.; Sumner, L. W.; Dixon, R. A. Profiling phenolic metabolites in transgenic alfalfa modified in lignin biosynthesis. *Phytochemistry* **2003**, *64*, 1013−1021.
(18) Deavours, B. E.; Dixon, R. A. Metabolic engineering of isoflavonoid biosynthesis in alfalfa (*Medicago sativa* L.). *Plant Physiol.* **2005**, *138*, 2245−2259.
(19) Aziz, N.; Paiva, N. L.; May, G. D.; Dixon, R. A. Profiling the transcriptome of alfalfa glandular trichomes. *Planta* **2005**, *221*, 28−38.
(20) Sumner, L. W.; Mendes, P.; Dixon, R. A. Plant metabolomics: large scale phytochemistry in the functional genomics era. *Phytochemistry* **2003**, *62*, 817−836.
(21) Mehrotra, B.; Mendes, P. Bioinformatics approaches to integrate metabolomics and other systems biology data. In *Biotechnology in Agriculture and Forestry, Vol. 57, Plant Metabolomics*; Saito, K., Dixon, R. A., Willmitzer, L., Eds.; Springer-Verlag: Berlin, Germany, 2006; pp 105−115.
(22) Rhee, S. Y.; Zhang, P.; Foerster, H.; Tissier, C. AraCyc: overview of an *Arabidopsis* metabolism database and its applications to plant research. In *Biotechnology in Agriculture and Forestry, Vol. 57, Plant Metabolomics*; Saito, K., Dixon, R. A., Willmitzer, L., Eds.; Springer-Verlag: Berlin, Germany, 2006; pp 141−153.
(23) Thimm, O.; Blaesing, O.; Gibon, Y.; Nagel, A.; Meyer, S.; Krueger, P.; Selbig, J.; Mueller, L.; Rhee, S. Y.; Stitt, M. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* **2004**, *37*, 914−939.
(24) Tokimatsu, T.; Sakurai, N.; Suzuki, H.; Shibata, D. KappA-View: a tool for integrating transcriptomic and metabolomic data on plant metabolic pathway maps. In *Biotechnology in Agriculture and Forestry, Vol. 57, Plant Metabolomics*; Saito, K., Dixon, R. A., Willmitzer, L., Eds.; Springer-Verlag: Berlin, Germany, 2006; pp 155−163.
(25) Iijima, Y.; Gang, D. R.; Lewinsohn, E.; Pichersky, E. Characterization of geraniol synthase from the peltate glands of sweet basil (*Ocimum basilicum*). *Plant Physiol.* **2004**, *134*, 370−379.
(26) Iijima, Y.; Davidovich-Rikanati, R.; Fridman, E.; Gang, D. R.; Bar, E.; Lewinsohn, E.; Pichersky, E. The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes in the peltate glands of three cultivars of basil (*Ocimum basilicum*). *Plant Physiol.* **2004**, *136*, 3724−3736.
(27) Jiang, H.; Xie, Z.; Koo, H.; McLaughlin, S. P.; Timmermann, B. N.; Gang, D. R. Metabolic profiling, phylogenetic analysis and anti-inflammatory investigation of *Zingiber* species: tools for authentication of ginger (*Zingiber officinale* Rosc.). *Phytochemistry* **2006**, *67*, 232−244 (doi: 10.1016/j.phytochem.2005.08.001).
(28) Jiang, H.; Sólyom, A.; Timmermann, B. N.; Gang, D. R. Characterization of gingerol-related compounds in ginger rhizome (*Zingiber officinale* Rosc.) by high-performance liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2957−2964.
(29) Jiang, H.; Somogyi, A.; Jacobsen, N. A.; Timmermann, B. N.; Gang, D. R. Analysis of curcuminoids by positive and negative electrospray ionization and tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1001−1012.
(30) Jiang, H.; Timmermann, B. N.; Gang, D. R. Use of liquid chromatography−electrospray ionization tandem mass spectrometry to identify diarylheptanoids in turmeric (*Curcuma longa* L.) rhizome. *J. Chromatogr. A* **2006**, *1111*, 21−31.
(31) Welham, T.; Domoney, C. Temporal and spatial activity of a promoter from a pea enzyme inhibitor gene and its exploitation for seed quality improvement. *Plant Sci.* **2000**, *159*, 289−299.
(32) Charlton, A.; Allnutt, T.; Holmes, S.; Chisholm, J.; Bean, S.; Ellis, N.; Mullineaux, P.; Oehlschlager, S. NMR profiling of transgenic peas. *Plant Biotechnol. J.* **2004**, *2*, 27−35.
(33) Charlton, A.; Godward, J.; Oehlschlager, S.; Ambrose, M.; Bowen, M.; Welham, T.; Mullineaux, P.; Domoney, C. Investigation of the metabolome of leaves from a pea germplasm array and from transgenic pea plants exposed to environmental stress. In *Proceedings of the 5th European Conference on Grain Legumes*, Dijon, France, 2004; pp 189−190.
(34) Jenkins, H.; Hardy, N.; Beckmann, M.; Draper, J.; Smith, A.; Taylor, J.; Fiehn, O.; Goodacre, R.; Bino, R.; Hall, R.; Kopka, J.; Lane, G.; Lange, B.; Liu, J.; Nikolau, B.; Mendes, P.; Oliver, S.; Paton, N.; Rhee, S.; Roessner-Tunali, U.; Saito, K.; Smedsgard, J.; Sumner, L.; Wang, T.; Walsh, S.; Wurtele, E.; Kell, D. B. A proposed framework for the description of plant metabolomics experiments and their results. *Nat. Biotechnol.* **2005**, *22*, 1601−1605.
(35) Fiehn, O.; Wohlgemuth, G.; Scholz, M. Automatic annotation of metabolomic mass spectra by integrating experimental metadata. *Proc. Lect. Notes Bioinformatics* **2005**, *3615*, 224−239.
(36) Catchpole, G. S.; Beckmann, M.; Enot, D. P.; Mondhe, M.; Zywicki, B.; Taylor, J.; Hardy, N.; Smith, A.; King, R. D.; Kell, D. B.; Fiehn, O.; Draper, J. Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14458−14462.
(37) Codex Alimentarius Commission. *Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants*; Codex Alimentarius Commission, 2003; CAC/GL 45-2003.
(38) FAO and WHO. *Safety Aspects of Genetically Modified Foods of Plant Origin*; report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology; WHO: Geneva, Switzerland, 2000.
(39) Kuiper, H. A.; Kleter, G. A.; Konig, A.; Hammes, W. P.; Knudsen, I. Concluding remarks special issue: safety assessment, detection and traceability, and societal aspects of genetically modified foods. European network on safety assessment of genetically modified food crops (ENTRANSFOOD). *Food Chem. Toxicol.* **2004**, *42*, 1047−1202.

(40) Cellini, F.; Chesson, A.; Colquhoun, I.; Constable, A.; Davies, H. V.; Engel, K. H.; Gatehouse, A. M. R.; Kärenlampi, S.; Kok, E. J.; Leguay, J.-J.; Lehesranta, S.; Noteborn, H. P. J. M.; Pedersen, J.; Smith, M. Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* **2004**, *42*, 1089−1125.

(41) Kuiper, H. A.; Kok, E. J.; Engel, K.-H. Exploitation of molecular profiling techniques for GM food safety assessment. *Curr. Opin. Biotechnol.* **2003**, *14*, 238−243.

(42) OECD Consensus Documents (http://www.oecd.org/document/9/0, 2340, en_2649_201185_1812041_1_1_1_1, 00.html).

(43) Ridley, W. P.; Sidhu, R. S.; Pyla, P. D.; Nemeth, M. A.; Breeze, M. L.; Astwood, J. D. A comparison of the nutritional profile of Roundup Ready corn event NK603 to that of conventional corn (*Zea mays* L.). *J. Agric. Food Chem.* **2002**, *50*, 7235−7243.

(44) George, C.; Ridley, W. P.; Obert, J. C.; Nemeth, M. A.; Breeze, M. L.; Astwood, J. D. Corn rootworm protected corn: composition of grain and forage from corn rootworm protected corn event MON 863 is equivalent to that of conventional corn (*Zea mays* L.), *J. Agric. Food Chem.* **2004**, *52*, 4149−4158.

(45) Reynolds, T. L.; Nemeth, M. A.; Glenn, K. C.; Ridley, W. P.; Astwood, J. D. Natural variability of metabolites in maize grain: differences due to genetic background. *J. Agric. Food Chem.* **2005**, *53*, 10061−10067.

(46) Tjeerdema, R. S.; Kauten, R. J.; Crosby, D. G. Sublethal effects of hypoxia in the abalone (*Haliotis rufescens*) as measured by in vivo ^{31}P NMR spectroscopy. *Comp. Biochem. Physiol.* **1991**, *100B*, 653−659.

(47) Tjeerdema, R. S. Using surface probe localized ^{31}P NMR spectroscopy to understand sublethal environmental actions. In *Multiple Stresses in Ecosystems*; Cech, J. J., Wilson, B. W., Crosby, D. G., Eds.; Lewis Publishers: Chelsea, MI, 1998; pp 155−180.

(48) Viant, M. R.; Pincetich, C. A.; Walton, J. H.; Tjeerdema, R. S.; Hinton, D. E. Utilizing in vivo NMR to study sublethal stress in aquatic organisms. *Mar. Environ. Res.* **2002**, *54*, 553−557.

(49) Pincetich, C. A.; Viant, M. R.; Hinton, D. E.; Tjeerdema, R. S. Metabolic changes in Japanese medaka (*Oryzias latipes*) during embryogenesis and hypoxia determined by in vivo ^{31}P NMR. *Comp. Biochem. Physiol.* **2005**, *140*, 103−113.

(50) Tjeerdema, R. S.; Fan, T. W.-M.; Higashi, R. M.; Crosby, D. G. Sublethal effects of pentachlorophenol in the abalone (*Haliotis rufescens*) as measured by in vivo ^{31}P NMR spectroscopy. *J. Biochem. Toxicol.* **1991**, *6*, 45−56.

(51) Viant, M. R.; Walton, J. H.; Tjeerdema, R. S. Comparative toxic actions of 3-trifluoro-4-nitrophenol (TFM) in marine mollsucs as characterized by in vivo ^{31}P-NMR. *Pestic. Biochem. Physiol.* **2001**, *71*, 40−47.

(52) Shofer, S. L.; Tjeerdema, R. S. Sublethal actions of pentachlorophenol in abalone (*Haliotis rufescens*) veliger larvae as measured by ^{31}P NMR. *Ecotoxicol. Environ. Saf.* **2002**, *51*, 155−160.

(53) Viant, M. R.; Walton, J. H.; TenBrook, P. L.; Tjeerdema, R. S. Sublethal actions of copper in abalone (*Haliotis rufescens*) as characterized by in vivo ^{31}P-NMR. *Aquat. Toxicol.* **2002**, *57*, 139−151.

(54) Tjeerdema, R. S.; Kauten, R. J.; Crosby, D. G. Interactive effects of pentachlorophenol and hypoxia in the abalone (*Haliotis rufescens*) as measured by in vivo ^{31}P NMR spectroscopy. *Aquat. Toxicol.* **1991**, *21*, 279−294.

(55) Tjeerdema, R. S.; Kauten, R. J.; Crosby, D. G. Interactive effects of pentachlorophenol and temperature in the abalone (*Haliotis rufescens*) as measured by in vivo ^{31}P NMR spectroscopy. *Aquat. Toxicol.* **1993**, *26*, 117−132.

(56) Shofer, S. L.; Willis, J. A.; Tjeerdema, R. S. Sublethal effects of pentachlorophenol and hypoxia on rates of arginine kinase flux in red abalone (*Haliotis rufescens*) as measured by ^{31}P magnetization saturation transfer NMR. *Mar. Environ. Res.* **1996**, *42*, 363−367.

(57) Tjeerdema, R. S.; Smith, W. S.; Martello, L. B.; Kauten, R. J.; Crosby, D. G. Interactions of chemical and natural stresses in the abalone (*Haliotis rufescens*) as measured by surface-probe localized ^{31}P NMR. *Mar. Environ. Res.* **1996**, *42*, 369−374.

(58) Shofer, S. L.; Willis, J. A.; Tjeerdema, R. S. Effects of hypoxia and toxicant exposure on arginine kinase function as measured by ^{31}P-NMR magnetization transfer in living abalone. *Comp. Biochem. Physiol.* **1997**, *117C*, 283−289.

(59) Shofer, S. L.; Willis, J. A.; Tjeerdema, R. S. Effects of hypoxia and toxicant exposure on phosphoarginine, intracellular pH, and free Mg^{2+} in abalone as measured by ^{31}P NMR. *Comp. Biochem. Physiol.* **1997**, *118A*, 1183−1191.

(60) Martello, L. B.; Tjeerdema, R. S.; Smith, W. S.; Kauten, R. J.; Crosby, D. G. Influence of salinity on the actions of pentachlorophenol in *Haliotis* as measured by in vivo ^{31}P NMR spectroscopy. *Aquat. Toxicol.* **1998**, *41*, 229−250.

(61) Viant, M. R.; Rosenblum, E. R.; Tjeerdema, R. S. NMR-based metabolomics: a powerful approach for characterizing the effects of environmental stressors on organism health. *Environ. Sci. Technol.* **2003**, *37*, 4982−4989.

(62) Viant, M. R.; Werner, I.; Rosenblum, E. R.; Gantner, A. S.; Tjeerdema, R. S.; Johnson, M. L. Correlation between heat-shock protein induction and reduced metabolic condition in juvenile steelhead trout (*Oncorhynchus mykiss*) chronically exposed to elevated temperature. *Fish Physiol. Biochem.* **2004**, *29*, 159−171.

(63) Rosenblum, E. S.; Viant, M. R.; Braid, B. M.; Moore, J. D.; Friedman, C. S.; Tjeerdema, R. S. Investigating the effects of pathogen, elevated temperature and starvation on the metabolic profiles of California red abalone, *Haliotis rufescens*. *Metabolomics* **2005**, *1*, 199−209.

(64) Viant, M. R.; Bundy, J. G.; Pincetich, C. A.; de Ropp, J.; Tjeerdema, R. S. NMR-derived developmental metabolic trajectories: an approach for visualizing the toxic actions of trichloroethylene during embryogenesis. *Metabolomics* **2005**, *1*, 149−158.

(65) Viant, M. R.; Pincetich, C. A.; Hinton, D. E.; Tjeerdema, R. S. Toxic effects of dinoseb in medaka (*Oryzias latipes*) embryos as determined by *in vivo* ^{31}P NMR, HPLC, and ^{1}H NMR metabolomics. *Aquat. Toxicol.* **2006**, *76*, 329−342.

(66) Viant, M. R.; Pincetich, C. A.; Tjeerdema, R. S. Metabolic effects of dinoseb, diazinon, and esfenvalerate in eyed eggs and alevins of Chinook salmon (*Oncorhynchus tshawytscha*) as determined by ^{1}H NMR metabolomics. *Aquat. Toxicol.* **2006**, in press.

(67) Glade, M. G. Food, nutrition, and the prevention of cancer: a global perspective. American Institute for Cancer Research/World Cancer Research Fund, American Institute for Cancer Research, 1997. *Nutrition* **1999**, *15*, 523−526.

(68) Verhoeven, D. T.; Goldbohm, R. A.; van Poppel, G.; Verhagen, H.; van den Brandt, P. A. Epidemiological studies on brassica vegetables and cancer risk. *Cancer Epidemiol. Biomarkers Prev.* **1996**, *5*, 733−748.

(69) Brown, A. F.; Yousef, G. G.; Jeffery, E. H.; Klein, B. P.; Kushad, M. M.; Wallig, M. A.; Juvik, J. A. Glucosinolate profiles in broccoli (*Brassica oleracea* L.): stability over environments and implications for cancer chemoprotection. *J. Am. Soc. Hortic. Sci.* **2002**, *127*, 807−813.

(70) Mithen, R.; Faulkner, K.; Magrath, R.; Rose, P.; Williamson, G.; Marquez, J. Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. *Theor. Appl. Genet.* **2003**, *106*, 727−734.

(71) Matusheski, N. V.; Wallig, M. A.; Juvik, J. A.; Klein, B. P.; Kushad, M. M.; Jeffery, E. H. Preparative HPLC method for the purification of sulforaphane and sulforaphane nitrile from *Brassica oleracea*. *J. Agric. Food Chem.* **2001**, *49*, 1867−1872.

(72) Matusheski, N. V.; Jeffery, E. H. A comparison of the bioactivity of two glucoraphanin hydrolysis products found in broccoli. *J. Agric. Food Chem.* **2001**, *49*, 5743−5749.

(73) Matusheski, N. V.; Swarup, R.; Juvik, J. A.; Mithen, R.; Bennett, M.; Jeffery, E. H. Epithiospecifier protein from broccoli (*Brassica oleracea* L. ssp. *italica*) inhibits formation of the anticancer agent sulforaphane. *J. Agric. Food Chem.* **2006**, *54*, 2069−2076.

(74) Foo, H. L.; Gronning, L. M.; Goodenough, L.; Bones, A. M.; Danielsen, B.; Whiting, D. A.; Rossiter, J. T. Purification and characterisation of epithiospecifier protein from *Brassica napus*: enzymic intramolecular sulphur addition within alkenyl thiohydroximates derived from alkenyl glucosinolate hydrolysis. *FEBS Lett.* **2000**, *468*, 243−246.

(75) German, J. B.; Bauman, D. E.; Burrin, D. G.; Failla, M. L.; Freake, H. C.; King, J. C.; Klein, S.; Milner, J. A.; Pelto, G. H.; Rasmussen, K. M.; Zeisel, S. H. Metabolomics in the opening decade of the 21st century: building the roads to individualized health. *J. Nutr*. **2004**, *134*, 2729−2732.

(76) Ward, R. E.; German, J. B. Understanding milk's bioactive components: a goal for the genomics toolbox. *J. Nutr*. **2004**, *134* (4), 962S−967S.

(77) Bhalla, R.; Narasimhan, K.; Swarup, S. Metabolomics and its role in understanding cellular responses in plants. *Plant Cell Rep.* **2005**, *24*, 562−571.

(78) Dunn, W. B.; Ellis, D. I. Metabolomics: current analytical platforms and methodologies. *Trends Anal. Chem.* **2005**, *24*, 285−294.

(79) Villa-Boas, S. G.; Rasmussen, S.; Lane, G. A. Metabolomics or metabolite profiles? *Trends Biotechnol.* **2005**, *23*, 385−386.

(80) Jenkins, H.; Johnson, H.; Kular, B.; Wang, T.; Hardy, N. Toward supportive data collection tools for plant metabolomics. *Plant Physiol.* **2005**, *138*, 67−77.

(81) Hirai, M. Y.; Klein, M.; Fujikawa, Y.; Yano, M.; Goodenowe, D.; Yamazaki, Y.; Kanaya, S.; Nakamura, Y.; Kitayama, M.; Suzuki, H.; Sakurai, N.; Shibata, D.; Tokuhisa, J.; Reichelt, M.; Gershenzon, J.; Papanbrock, J.; Saito, K. Elucidation of gene-to-gene and metabolite-to-gene networks in *Arabidopsis* by integration of metabolomics and transcriptomics. *J. Biol. Chem.* **2005**, 25590−25595.

(82) Goodacre, R. Making sense of the metabolome using evolutionary computation: seeing the wood with the trees. *J. Exp. Bot.* **2005**, *56*, 245−254.

(83) Tokimatsu, T.; Sakurai, N.; Suzuki, H.; Ohta, H.; Nishitani, K.; Koyama, T.; Umezawa, T.; Misawa, N.; Saito, K.; Shibata, D. KaPPA-view. A web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiol.* **2005**, 1289−1300.

Received for review May 1, 2006. Revised manuscript received August 18, 2006. Accepted August 19, 2006.

JF061218T